Econometria
Propiedades Estadísticas Modelo de Regresión Lineal Simple

Pasquini, Ricardo

IAE
Universidad Austral

16 de septiembre de 2024

Modelo Econométrico

- Definición del modelo econometrico con un solo regresor
- Notación: $Y = \beta_0 + \beta_1 X + \varepsilon$
- Explicación de términos: Y (variable dependiente), X (variable independiente), β_0 (intercepto), β_1 (coeficiente de pendiente), y ε (término de error)

Estimador de Mínimos Cuadrados Ordinarios (OLS)

- Definición del estimador OLS: $\hat{\beta}_1$
- El estimador OLS minimiza la suma de los residuos al cuadrado
- Fórmula: $\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$

Ausencia de Sesgo del Estimador OLS

- Definición de estimador sin sesgo
- El estimador OLS no tiene sesgo:

$$E(\hat{\beta}_1) = \beta_1$$

- Comprensión intuitiva de la ausencia de sesgo: En promedio, el estimador OLS es igual al verdadero parámetro de población

Supuestos de OLS para garantizar que no hay sesgo

- Supuestos necesarios para la ausencia de sesgo del estimador OLS:
  1. Linealidad en los parámetros (el modelo poblacional es $y = \beta_0 + \beta_1 X + \varepsilon$)
  2. Muestreo aleatorio (los valores (x_i, y_i) son variables aleatorias del modelo poblacional)
  3. La esperanza condicional del error es cero ($E(\varepsilon|X) = 0$)

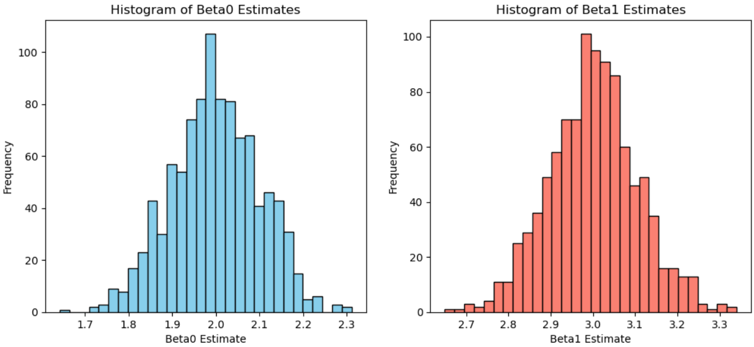
Ausencia de Sesgo del Estimador OLS

- Usando los supuestos, se puede demotrar el resultado.
- Es util hacerlo en dos pasos.
- Primero mostrar que $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \bar{x})\varepsilon_i}{SST_x}$
- Luego $E[\hat{\beta}_1|x] = \beta_1 + \frac{1}{SST_x}\sum(x_i - \bar{x})E[\varepsilon_i|x] = \beta_1$

Idea de la Simulación

- **Objetivo**: Mostrar que los coeficientes del modelo OLS son insesgados.
- **Pasos de la simulación**:
  - **Definición de parámetros**: Establecemos los valores verdaderos de los coeficientes asumidos verdaderos (β_0 y β_1), tamaño de muestra y número de simulaciones.
  - **Iteración**: Realizar múltiples simulaciones.
    - Generamos pares (x,y) provenientes de la poblacion (cumplen $y = \beta_0 + \beta_1 X + \epsilon$)
    - Ajustamos el modelo de regresión lineal y obtenemos los estimadores de los coeficientes.
    - Almacenar los valores estimados de los coeficientes.
  - **Visualización**: Construimos histogramas de los valores estimados de los coeficientes.
  - **Estadísticas descriptivas**: Calculamos media y desviación estándar de los valores estimados de los coeficientes.

Ausencia de Sesgo del Estimador OLS



Varianza del Estimador OLS

- Denotamos la varianza del estimador OLS como $Var(\hat{\beta}_1)$
- Se puede derivar que la varianza del estimador OLS esta dada por

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$$

- Donde $\sigma^2 = Var(\varepsilon)$
- Interpretación: mide la dispersión o variabilidad de la distribución muestral del estimador OLS.
- Notar que la varianza del estimador es mayor a mayor error y menor a mayor variabildad del x!

Varianza del Estimador OLS - Derivación

- Para derivar este resultado partimos de $\hat{\beta}_1 = \beta_1 + \frac{\sum(x_i - \bar{x})\varepsilon_i}{SST_x}$
- Luego $Var[\hat{\beta}_1|x] = \frac{1}{SST_x^2}\sum(x_i - \bar{x})^2 Var(\varepsilon_i)$

Conclusión

- La ausencia de sesgo del estimador es una propiedad deseable pero no la que necesariamente quisieramos tener en todos los casos.
- Entender los determinantes de la varianza es crítico para entender la confiabilidad de nuestra estimación, como veremos al realizar tests de hipótesis.
- Simular es una herramienta importante e imprescindible al hacer estadística/econometría.

Referencias

- Wooldridge ch2